Optimal Network Estimation of Origin-Destination Flow from Sufficient Link Data

Fabien Leurent, Frédéric Meunier

PAPER SUBMITTED FOR 88th TRB MEETING, 2009, AND FOR PUBLICATION IN TRB JOURNAL

TRB PAPER # 09-1885

PAPER TITLE

Optimal Network Estimation of Origin-Destination Flow from Sufficient Link Data

AUTHORS' NAMES

Fabien LEURENT, Frédéric MEUNIER

SUBMISSION DATE

July 31, 2008

Revised version: December 1, 2008

AUTHOR 1 (CORRESPONDING AUTHOR)

Prof. Fabien Leurent

Université Paris-Est, Laboratoire Ville Mobilité Transport

19 rue Alfred Nobel, Champs sur Marne, 77455 Marne la Vallée Cedex 2, France

Tel: (33) 164 152 111

Fax: (33) 164 152 140

e-mail: fabien.leurent@enpc.fr

AUTHOR 2

Frédéric Meunier

Université Paris-Est, Laboratoire Ville Mobilité Transport

19 rue Alfred Nobel, Champs sur Marne, 77455 Marne la Vallée Cedex 2, France

Tel: (33) 164 152 121

Fax: (33) 164 152 140

e-mail: frederic.meunier@enpc.fr

WORD COUNT: 6,000

PLUS 4 FIGURES AND 2 TABLES

Optimal Network Estimation of Origin-Destination Flow from Link Data

Fabien Leurent, Frédéric Meunier (Université Paris-Est, Lvmt)

Abstract

A systematic method is provided to estimate an origin destination flow at the network level on the basis of link traffic counts with OD structure. The estimator has minimal variance among the unbiased linear combinations of link estimators. The problem to find the optimal estimator is stated as a linear system in node potentials; it is endowed with nice properties of existence and uniqueness, which enjoy a structural, graph-based interpretation. Assuming independent local estimators, an equivalent, dual problem is that of minimizing the imprecision energy carried by a network flow.

The estimation method does not require assumptions on route choice proportions. Throughout the paper, a realistic case is addressed both to discuss the practical issues and to demonstrate the solution method.

Keywords

Graph cut. Graph Laplacian. Feasible differential problem. Imprecision flow. Unbiased estimator. Variance minimization.

Manuscript Text

1. INTRODUCTION

As a transportation network is purported to carry trips from origin points to destination points, the pattern of flow by origin-destination (OD) pair is a major issue both at the local level of the network link in order to reveal its spatial function as a transport element, and at the global level of the whole network in order to indicate which links and routes do accommodate specific OD flows. Knowing the network pattern of OD flows enables a planner to design route signalling and other network operation plans, as well as to evaluate the adequacy of network schemes to the needs of the transport users. Specifically, the matrix of OD flows (by time period and user class) is a basic input in network assignment models which are used to simulate transport schemes and to evaluate their costs and benefits (*1*, *2*, *3*).

This has motivated the development of specific surveys to obtain OD information (*4*): from household- or firm-based interview surveys that describe the trips made by the interviewee over a given period, to en route surveys in which the trip-maker is asked about his origin and destination points, passing by "cordon" surveys in which a given trip-maker is identified at a couple of transit points. All of the network-based surveys yield local information, which implies that a system of local surveys is required to describe the pattern of OD flows at the network level.

The objective of this paper is to bring about an optimal statistical estimator of the trip flow of an OD pair at the network level, on the basis of link-based OD flow estimators which by assumption are unbiased, of given accuracy and independent. The network estimator is unbiased and has minimal variance among the unbiased estimators that are linear combinations of the local estimators. Owing to a graph-theoretic analysis, the optimal linear combination is characterized as the solution of a quadratic minimization program subject to linear constraints. This is solved in an efficient way, yielding easy graph-theoretic formulae for the mean and variance values of the optimal estimator of the OD flow at the network level. A basic requirement for existence is that any path from origin to destination must traverse along at least one of the links with local information. As the solution method is easy to implement on a computer, our methodology would be useful in any planning study in which an OD matrix has to be recovered from "relatively abundant" local information – meaning more than sufficient: this case arises notably in interurban roadway traffic studies.

The remainder of the paper is organized in four parts. Section 2 provides a statistical analysis of link-based surveys and their combination to estimate an OD flow at the network level, in a bottom-up approach. The by-hand computation is shown to be involved, which demonstrates the scope for a systematic, computer-based procedure. In Section 3, the problem of optimal estimation is worked out and endowed with some graph-theoretic properties of characterization, existence and uniqueness. It is shown to be equivalent to a network electrical problem with link resistance equal to the link accuracy (of the local estimator of the link OD flow). The solution method, studied in Section 4, is a straightforward approach to solve the characteristic property of Section 3 by computing specific node potentials: it is appropriate and efficient when all links are informative, and also when some links are uninformative (no data hence infinite variance) on the basis of graph contraction. Lastly, Section 5 provides concluding comments, together with some research topics.

2. ESTIMATION OF OD FLOW FROM LINK INFORMATION

In this Section, a bottom-up approach is taken to conduct the statistical analysis from the local level of a link to the global level of the network. At the link level, the main issue is to obtain a statistical estimator of the link OD flow: the usual method is to combine an estimator of link flow (Subsection 2.1) to an estimator of the proportion of OD pair within link flow (Subsection 2.2), in a multiplicative way (Subsection 2.3). At the network level, the link estimators must be combined in parallel along an "informative graph cut" between the origin and destination nodes (Subsection 2.4). Furthermore, when several informative cuts are available for a given OD pair, it is more efficient to make use of all the informative links: this is accomplished through Bayesian mix (Subsection 2.5).

Our bottom-up approach will be illustrated using a French instance of interurban road traffic: in the Poitou-Charentes region located in the midway of Paris and Bordeaux, link OD surveys were performed on a set of main roads by a consultancy firm, the Cété Sud-Ouest: see Figure 1. Taking the cities of Angoulême as origin and Rochefort as destination, there are five informative links denoted by A to E and depicted in Figure 2. The graph cut {A, B} is an informative cut for the OD pair; so is the {B, C, D, E} cut. Tables 1 and 2 provide the associated data, to be explained in the course of the section. The material in Section 2 is taken from (*5*).

FIGURE 1 Main road network and location of link surveys (black triangles)

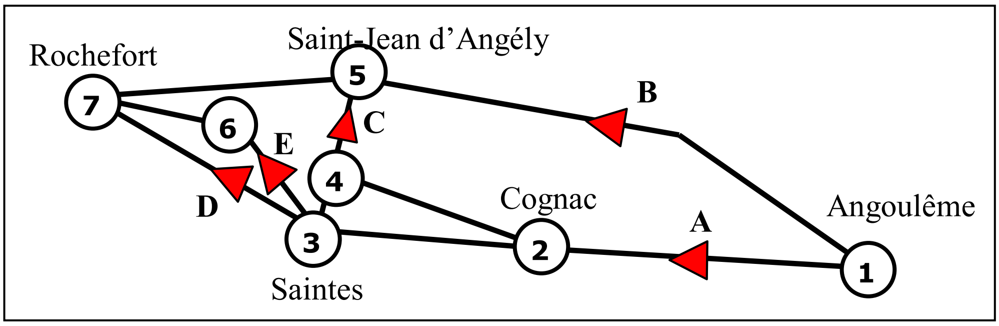

FIGURE 2. Informative links for OD pair Angoulême-Ro

TABLE 1 Link Count Data

Arc *a*	Sample size, # days of count, N_a	Sample mean of daily count $\hat{q}_a$	Sample corrected standard deviation $S^*[(Q_{an})_n]$	Standard error $SE[\hat{q}_a]$
A	15	10,029	3,824	987.4
B	5	3,739	1,260	563.5
C	10	7,107	2,720	860.1
D	20	9,735	3,645	815.0
E	12	5,736	3,400	981.5

TABLE 2 Link OD Survey Data

	Sample size	# cases of OD i	Sample mean	Standard error	Link OD flow	
Arc *a*	# cases, n_a	n_{ia}	$\hat{p}_{ia}$	$SE[\hat{p}_{ia}]$	$\hat{q}_{ia}$	$SE[\hat{q}_{ia}]$
A	1,332	11	0.83%	0.25%	82.8	26.3
B	676	13	1.92%	0.53%	71.9	22.7
C	1,243	0	0.00%	0.00%	0.0	0,0
D	1,388	14	1.01%	0.27%	98.2	27.5
E	1,554	14	0.90%	0.24%	51.4	16.5

2.1 The estimation of link flow

Let us denote by a an arc, which is an oriented link to be used by a trip only in the direction from its tail node to its head node. Let A be the set of network arcs. Denote by I the set of OD pairs: in fact an OD pair i is a couple of two zones that accommodate the origin [resp. destination] endpoints of trips.

The issue is to estimate a link and OD (LOD) flow q_{ia}, which stands for the mean flow over a set of homogeneous time periods, for instance the days in a year. A classical estimator $\hat{q}_{ia}$ of q_{ia} is defined as the product of two terms, first an estimator $\hat{q}_a$ of the link flow q_a and second an estimator $\hat{p}_{ia}$ of the OD proportion p_{ia} along that link. An estimator $\hat{q}_a$ is usually derived by taking the (unweighted) mean of N_a sampled period flows $[Q_{an}]_{n=1,..N_a}$: it holds that

$$\hat{q}_a = \frac{1}{N_a}\sum_{n=1}^{N_a} Q_{an} .$$

Assuming that the sampling scheme is unbiased, the link flow estimator is unbiased, and over the population of samples:

$$\mathrm{E}[\hat{q}_a] = \frac{1}{N_a}\sum_{n=1}^{N_a} \mathrm{E}[Q_{an}] = \mathrm{E}[Q_{an}] = q_a .$$

The sample is also useful to estimate the variance of the link flow over the distribution of periods, by means of the following formula of corrected sample variance

$$S_{q_a}^{2*} = \frac{1}{N_a - 1}\sum_{n=1}^{N_a} [q_{an} - \hat{q}_a]^2 .$$

Assuming unbiased, independent sampling it holds that the sample mean has a standard deviation of

$$\sigma_{\hat{q}_a} = \sqrt{\mathrm{var}[\hat{q}_a]} = \sigma_{q_a} / \sqrt{N_a} .$$

This is approximated without bias by the sample standard error, namely

$$\mathrm{SE}[\hat{q}_a] = \sqrt{S_{q_a}^{2*}} / \sqrt{N_a} \approx \sigma_{\hat{q}_a} .$$

Instance. In Table 1 the mean flow on link A is estimated as 10,029 veh/day, with standard error of 987.4 veh/day.

2.2 The estimation of OD proportion within link flow

The next task is to estimate the proportion p_{ia} of OD pair i within link flow q_a. A classical estimator $\hat{p}_{ia}$ of p_{ia} is derived from an en route OD survey, in which a sample of trips is taken so as to reveal their origin and destination points. Under the no bias assumption, out of n_a sampled trips there are n_{ia} trips which belong to OD pair i, yielding the following estimator of empirical frequency:

$$\hat{p}_{ia} = \frac{n_{ia}}{n_a} .$$

This is because the binary variable "Whether or Not to belong to OD pair i", $y_{ia} \in \{1,0\}$ for the trips through link *a*, has mean p_{ia} and sample mean of $\hat{p}_{ia} = n_{ia} / n_a$ since $n_{ia} = \sum_{n=1}^{n_a} y_{ian}$. Then, over the population of samples,

$$\mathrm{E}[\hat{p}_{ia}] = p_{ia}$$

Moreover the binary variable has distribution variance $p_{ia}(1 - p_{ia})$. Over the population of samples the random variable $\hat{p}_{ia}$ has variance $\mathrm{var}[n_{ia}] / n_a^2 = \sigma_{y_{ia}}^2 / n_a$, hence a standard deviation of $\sigma_{y_{ia}} / \sqrt{n_a}$. Under the independence assumption, the uncorrected sample variance $S_{y_{ia}}^2 = (\sum_{n=1}^{n_a}[y_{ian} - \hat{p}_{ia}]^2)/n_a$ reduces to $\hat{p}_{ia}(1-\hat{p}_{ia})$; the standard error on $\hat{p}_{ia}$ is

$$\mathrm{SE}[\hat{p}_{ia}] = \sqrt{\frac{\hat{p}_{ia}(1-\hat{p}_{ia})}{n_a - 1}} \approx \sigma_{\hat{\pi}_{ia}}.$$

Instance. In Table 1 the proportion of the OD pair Angoulême-Rochefort in the traffic of link A is estimated as $\hat{p}_{ia} =$ 0.83% with standard error of 0.25%.

2.3 Derivation of Link OD estimator

The last task at the local level is to multiply the estimator of link flow $\hat{q}_a$ by that of OD proportion $\hat{p}_{ia}$ to yield the estimator $\hat{q}_{ia}$ of the link and origin-destination (LOD) flow $q_{ia} = q_a p_{ia}$, namely

$$\hat{q}_{ia} = \hat{q}_a \hat{p}_{ia}.$$

Assuming that both basic estimators are unbiased and that they are independent of each other, their product is an unbiased estimator of q_{ia}, since:

$$\mathrm{E}[\hat{q}_{ia}] = \mathrm{E}[\hat{q}_a]\mathrm{E}[\hat{p}_{ia}] = q_a p_{ia} = q_{ia}.$$

Furthermore, as for any two independent random variables X and Y their product has variance $\mathrm{var}[XY] = \mathrm{var}[X]\mathrm{var}[Y] + \mathrm{E}[X]^2\,\mathrm{var}[Y] + \mathrm{E}[Y]^2\,\mathrm{var}[X]$ ([1]), it holds that

$$\mathrm{var}[\hat{q}_{ia}] = \mathrm{var}[\hat{q}_a]\mathrm{var}[\hat{p}_{ia}] + \mathrm{E}[\hat{q}_a]^2\,\mathrm{var}[\hat{p}_{ia}] + \mathrm{E}[\hat{p}_{ia}]^2\,\mathrm{var}[\hat{q}_a]$$

Hence $\mathrm{SE}^2[\hat{q}_{ia}] \approx \mathrm{SE}^2[\hat{q}_a]\,\mathrm{SE}^2[\hat{p}_{ia}] + \hat{q}_a^2\,\mathrm{SE}^2[\hat{p}_{ia}] + \hat{p}_{ia}^2\,\mathrm{SE}^2[\hat{q}_a]$.

Instance. The flow of OD pair Angoulême-Rochefort through link A is estimated by $\hat{q}_{i/\mathrm{A}} =$ 82.8 veh/day, with $\mathrm{SE}[\hat{q}_{i/\mathrm{A}}] =$ 26.3 veh/day.

The material in Subsection 2.1 and 2.2 is standard statistical background, see e.g. (*6*). However the product of random variables is a less well-known topic apart from the mean formula when X and Y are independent, $\mathrm{E}[XY] = \mathrm{E}[X]\mathrm{E}[Y]$. The variance formula was applied to an LOD flow in (*7*).

2.4 On parallel combination and graph cut

Let us now turn our attention to a network OD flow: as any trip is made along a network route, the OD flow is taken as the total route flow over all the routes that take from the origin to the destination. Moreover, since link OD flows rather than route flows are available, let us search for a subset of links that intercept all the OD routes in a precise way, each one with an interception coefficient of +1: then the total link flow over that subset makes an estimator of the OD flow. In graph theory, such a subset is called a graph cut $j = [S : N \setminus S]$ separating the destination node from the origin node: it is made up of the "parallel" arcs that separate a node subset S which includes o from the

[1] $\mathrm{var}[XY] = \mathrm{E}[X^2Y^2] - \mathrm{E}[XY]^2 = \mathrm{E}[X^2]\mathrm{E}[Y^2] - \mathrm{E}[X]^2\mathrm{E}[Y]^2 = \mathrm{E}[X^2](\mathrm{E}[Y^2] - \mathrm{E}[Y]^2)$ $+ \mathrm{E}[Y]^2(\mathrm{E}[X^2] - \mathrm{E}[X]^2)$ to be averaged with the symmetrical statement of $\mathrm{var}[YX]$

complementary node subset $N \setminus S$ which includes d: every $o-d$ route r traverses j at least once, and the number of oriented traversals of j by r amounts to +1, thus satisfying the requirement on the interception coefficient.

An arc a is in j either with positive orientation denoted by $e_j(a) := +1$ if its tail is in S and its head in $N \setminus S$ or with negative orientation denoted by $e_j(a) := -1$ if its tail is in $N \setminus S$ and its head in S. Let also $e_j(a) := 0$ for $a \notin j$. Along cut j, the network OD flow is estimated as

$$\hat{q}_{i/j} = \sum_{a \in j} e_j(a) \hat{q}_{ia} .$$

Assuming unbiased OD flow estimators, this is an unbiased estimator since

$$\mathrm{E}[\hat{q}_{i/j}] = \sum_{a \in j} e_j(a) \mathrm{E}[\hat{q}_{ia}] = \sum_{a \in j} e_j(a) q_{ia} = q_i .$$

Assuming further that the information sources are independent, then

$$\mathrm{var}[\hat{q}_{i/j}] = \sum_{a \in j} \mathrm{var}[\hat{q}_{ia}] \text{, hence}$$

$$\mathrm{SE}[\hat{q}_{i/j}] = \sqrt{\sum_{a \in j} \mathrm{SE}^2[\hat{q}_{ia}]} .$$

Instance. Let us come back to our example. The subset {A, B} makes a first cut from Angoulême to Rochefort, with $\hat{q}_{i/\mathrm{A,B}} =$ 154.7 veh/day and $\mathrm{SE}[\hat{q}_{i/\mathrm{A,B}}] =$ 34.7 veh/day. A second cut $\{\mathrm{B, C, D, E}\}$ provides an estimation of $\hat{q}_{i/\mathrm{B,C,D,E}} =$ 251.5 veh/day and $\mathrm{SE}^2[\hat{q}_{i/\mathrm{B,C,D,E}}] =$ 39.3 veh/day.

Out of the alternative informative cuts, which one should be preferred for our purpose of estimation? The intuitive answer is to select the cut of minimal variance hence of maximal accuracy. In our example, the first cut is better than the second one since $\mathrm{SE}[\hat{q}_{i/\mathrm{A,B}}] < \mathrm{SE}[\hat{q}_{i/\mathrm{B,C,D,E}}]$.

2.5 Bayesian mixture of redundant information sources

The availability of alternative informative cuts leads us to the main objective in this paper: how could we combine various local, link-based OD flow estimators in an efficient way, and eventually in an optimal way?

Considering again our example, the estimator of link A is redundant with the parallel combination of links C, D, E: such a redundancy should not induce us to reject the less informative source, but rather to combine both sources so as to maximize the resulting accuracy.

The relevant statistical tool is called the Bayesian mixture of information sources: Bayesian analysis is purported to combine actual observation with prior information in order to yield posterior information. For any two real independent random variables X and Y that describe a unique phenomenon, the Bayesian mix $X \otimes Y$ is a random variable with variance such that

$$\frac{1}{\mathrm{var}[X \otimes Y]} = \frac{1}{\mathrm{var}[X]} + \frac{1}{\mathrm{var}[Y]}$$

Put another way, the accuracy of the mix is the sum of those of the basic variables. Furthermore,

$$\mathrm{E}[X \otimes Y] = \frac{\mathrm{var}[X \otimes Y]}{\mathrm{var}[X]}\mathrm{E}[X] + \frac{\mathrm{var}[X \otimes Y]}{\mathrm{var}[Y]}\mathrm{E}[Y],$$

which means that the mix mean is a convex combination of the basic means weighted by their relative share in the mix accuracy. More elaborate formulae exist for multidimensional variables (denoting by C_X the matrix of variance-covariance of random vector X, the mixture $Z = X \otimes Y$ satisfies $C_Z^{-1} = C_X^{-1} + C_Y^{-1}$ and $E_Z = C_Z[C_X^{-1}E_X + C_Y^{-1}E_Y]$).

An instance of application pertains to sampling schemes with respect to sample size for a variable *X*. Mixing a sample of size *N* with another sample of size *M* yields a sample of size $N + M$. The variances of sample means satisfy

$$\mathrm{var}[\hat{X}_N] = \mathrm{var}[X]/N,$$

$$\mathrm{var}[\hat{X}_M] = \mathrm{var}[X]/M,$$

$$\mathrm{var}[\hat{X}_{N+M}] = \mathrm{var}[X]/(N+M),$$

which corresponds to the formula of the mix variance, since $\hat{X}_{N+M} = \hat{X}_N \otimes \hat{X}_M$. The means of the samples also comply with the formula of the mix mean, since

$$\begin{aligned}\mathrm{E}[\hat{X}_{N+M}] &= \mathrm{E}[X] \\ &= \tfrac{N}{N+M}\mathrm{E}[\hat{X}_N] + \tfrac{M}{N+M}\mathrm{E}[\hat{X}_M]\end{aligned}.$$

For our purpose of OD flow estimation, let α and β denote two link subsets that bring redundant information $\hat{q}_\alpha, \hat{q}_\beta$: then

$$\mathrm{SE}^{-2}[\hat{q}_\alpha \otimes \hat{q}_\beta] = \mathrm{SE}^{-2}[\hat{q}_\alpha] + \mathrm{SE}^{-2}[\hat{q}_\beta]$$

$$\mathrm{E}[\hat{q}_\alpha \otimes \hat{q}_\beta] = \frac{\mathrm{SE}^2[\hat{q}_\alpha \otimes \hat{q}_\beta]}{\mathrm{SE}^2[\hat{q}_\alpha]}\mathrm{E}[\hat{q}_\alpha] + \frac{\mathrm{SE}^2[\hat{q}_\alpha \otimes \hat{q}_\beta]}{\mathrm{SE}^2[\hat{q}_\beta]}\mathrm{E}[\hat{q}_\beta]$$

Instance. Coming back to our example, by parallel combination the pooling of links C, D, E yields (partial) OD flow of mean value 149.6 veh/day and standard error of 32.0 veh/day. Mixing that with the OD flow on link A yields partial OD flow of mean value 109.7 veh/day and standard error of 20.3 veh/day. The parallel combination of the pool with link B yields an overall OD flow of mean value 181.6 veh/day and standard error of 30.5 veh/day. This makes a significant improvement over the standard error of any informative cut, and provides an intermediary estimate of the mean OD flow, which is valuable because of the large discrepancy between the alternative cut estimates.

2.6 Comments

Two assumptions of independence play a crucial role in the analysis. First, at the link level we require the link count and the OD survey to be independent: indeed this is a strong assumption, especially so in interurban road traffic where there is considerable day-to-day variation, whereas for economical stakes the OD survey is focused on one or a few days. Then the sample of OD pairs does represent the traffic of the surveyed period(s), rather than the whole population of trips across the set of periods. This is an issue of ill-representation or of indirect dependency, rather than of statistical dependency since the statistical protocols for link counts on one hand, link OD survey on the other hand, may perfectly well be independent; however the time restriction of the OD survey induces an indirect dependency through the specific structure of traffic in the surveyed period, as compared to the overall structure of traffic across the set of periods.

Second, in the framework of Bayesian mixture the basic variables are required to be independent: thus in the example we were driven to take B apart and identify that A is redundant with {C, D, E}, instead of pooling {A, B} with {C, D, E, B} in a straightforward way. In practice, a straightforward pooling of the informative cuts would of course yield improvement over the selection of only one informative cut, be it that of least variance. Taking the mean of the two cuts would yield a mean OD flow of 203.1 veh/day – also an intermediary value – and an approximate (yet flawed) standard error of 26.2 veh/day.

Therefore the main issue is to design a systematic procedure to combine the local information sources into a network estimator of OD flow. Admittedly, although our example was designed as a classroom case, its by-hand treatment was somewhat involved and tedious. The next Sections provide a systematic procedure, of which the computer implementation is easy.

3. OPTIMAL ESTIMATOR WITH GRAPHICAL PROPERTIES

Let us now state the problem of optimal estimation and provide a mathematical analysis and characterization. In Subsection 3.1 some notation is given and some graph-theoretic properties are recalled. Then, in Subsection 3.2 the problem of optimal estimation is stated on the set of linear combinations of link OD flows, as an unbiased combination of minimal variance: two lemmas are helpful, first to characterize an unbiased linear combination and second to relate it to a network field of node potentials. Next, Subsection 3.3 is devoted to the derivation of optimality conditions, which amount to a linear system in the node potentials. In Subsection 3.4, the issues of existence and uniqueness of an optimal estimator are addressed. Subsection 3.5 provides an alternative, equivalent form as an electrical network problem of energy minimization, where energy pertains to statistical imprecision. Lastly, the framework is adapted to relax the assumption of independence between the local estimators.

3.1 Notation and assumptions

We consider a network graph $G=[N,A]$ where N is the set of *nodes n* and A the set of *arcs a* with tail node n_a^+ and head node n_a^-. Two particular nodes are given: o, the *origin*, and d, the *destination*. A network *route r* is a sequence of distinct arcs $a_1, a_2, ..., a_\ell$ such that for any i the head of a_i is the tail of a_{i+1}.

Let us recall the definition of the node-arc incidence functions:

- positive incidence function: $e^+(n,a)=1$ if *a* leaves *n* or 0 otherwise,
- negative incidence function: $e^-(n,a)=1$ if *a* enters *n* or 0 otherwise,
- (net) incidence function: $e(n,a)=e^+(n,a)-e^-(n,a)$.

The *incidence matrix* E is the $N \times A$ matrix whose entries are the $e(n,a)$. If the graph $G=[N,A]$ has weights w_a on its arcs, the *weighted incidence matrix* E_w is the $N \times A$ matrix whose entries are the $e(n,a)w_a$.

The *Laplacian matrix* L_w of a multigraph $G=[N,A]$ with arc weights w_a is a symmetric, $N \times N$ matrix defined by

$$\mathrm{L}_w(n,n') = \begin{cases} \sum_{a:e(n,a)\neq 0} w_a & \text{if } n=n' \\ 0 & \text{if } n \neq n' \text{ and } n \text{ and } n' \text{ are non-adjacent} \\ \sum_{a:e(n,a)e(n',a)\neq 0} -w_a & \text{if } (n,n') \text{ or } (n',n) \text{ is an arc } a \end{cases}$$

The last line means that if n and n' are adjacent, the value of the corresponding entry in the Laplacian matrix is the sum of the weights of all arcs having n and n' as endpoints. In particular, if G is a simple graph, this sum has at most one term: $w_{(n,n')}$ or $w_{(n',n)}$ (whether (n,n') or (n',n) exists).

The following property will be useful to us:

$$\mathrm{E}_w \mathrm{E}_w^t = \mathrm{L}_{w^2}.$$

The OD pair $i=(o,d)$ is serviced by o-d routes $r \in R_i$ such that any arc a cannot be traversed more than once by *r*: then the arc-route incidence function $e_r(a)$ is defined by $e_r(a)=1$ if $a \in r$ or $e_r(a)=0$ if $a \notin r$. The OD flow q_i is made up of the route flows

$$q_i = \sum_{r \in R_i} q_r = \sum_{r \in i} q_r.$$

Our objective is to recover the mean OD flow $\overline{q}_i$ of the underlying random variable q_i, which induces random flows q_r of mean value $\overline{q}_r$. The information sources are link OD flow estimators $\hat{q}_{ia}$, which we assume to have no bias and variance σ_a^2. Furthermore, the OD flow estimators of distinct links are assumed to be independent hence $\mathrm{cov}[\hat{q}_{ia}, \hat{q}_{ib}] = 0$ if $a \neq b$.

3.2 Problem statement and solution

Let us now state the problem of optimal estimator in a precise way as follows:

"Given a set of unbiased, independent link OD flow estimators $\hat{q}_{ia}$ of variance σ_a^2, find a network estimator $\hat{q}_i = \sum_{a \in A} \lambda_a \hat{q}_{ia}$ as a linear combination of the link estimators with real coefficients λ_a, such that it has no bias i.e. $\mathrm{E}[\hat{q}_i] = \overline{q}_i$ and minimal variance $\mathrm{var}[\hat{q}_i] = \sum_{a \in A} \sigma_a^2 \lambda_a^2$ among the set of unbiased linear combinations".

Our mathematical analysis is aimed to demonstrate the theorem that follows.

Theorem 1, Optimal linear unbiased estimator for network OD flow. *The linear estimator $\hat{q}_i := \sum_{a \in A} \lambda_a \hat{q}_{ia}$ of minimal variance is such that $\lambda_a = \pi(\mathrm{n}_a^-) - \pi(\mathrm{n}_a^+)$ for every arc $a \approx (\mathrm{n}_a^+, \mathrm{n}_a^-) \in A$, where the potential field $\pi \in \Re^N$ solves the following system*

$$(\mathbf{L}_{\sigma^2}\pi)(n) = 0 \ \text{ for } n \notin \{o,d\}$$
$$\pi(o) = 0 \tag{1}$$
$$\pi(d) = 1$$

The system above has $|N|$ linear constraints and $|N|$ unknowns. Hence, it is easy to compute the linear, unbiased estimator of optimal variance (for instance on the basis of Gaussian elimination). When $\hat{q}_{ia}$ is not available for some arcs a, it is still possible to compute the best linear estimator with a simple technique of graph contraction that is explained in the next Section (provided that there be at least one $o-d$ cut with $\hat{q}_{ia}$ available on every arc).

The proof of Theorem 1 is based on two lemmas: first, to characterize an unbiased linear estimator by a set of conditions on the coefficients $(\lambda_a)_{a\in A}$ along any $o-d$ route; second, to replace these conditions about the routes by an equivalent condition which pertains to node potentials and is much more tractable.

Lemma 1, Unbiased linear estimator. *The linear estimator* $\hat{q}_i = \sum_{a\in A}\lambda_a\hat{q}_{ia}$ *is an unbiased estimator of the* $o-d$ *flow if and only if one has* $\sum_{a\in r}\lambda_a = 1$ *for all* $o-d$ *routes* $r \in R_i$.

Proof. The random variable $\hat{q}_i$ is an unbiased estimator if and only if

$$\mathrm{E}[\hat{q}_i] = \overline{q}_i = \sum_{r\in i}\overline{q}_r \ . \tag{2}$$

First, $\hat{q}_i = \sum_{a\in A}\lambda_a\hat{q}_{ia}$. Second, $\hat{q}_{ia} = \sum_{r\in i}\hat{q}_{r/a}$ in which the flow of route r as intercepted on arc a is indexed also by "$/a$" to indicate that the estimator is notional (it is not required to have it, we only use the link OD estimator): as we assume unbiased link OD flow, consistently the notional estimator should be unbiased hence $\mathrm{E}[\hat{q}_{r/a}] = \overline{q}_r$. On combining, we get that $\mathrm{E}[\hat{q}_i] = \sum_{a\in A}\lambda_a\sum_{r\in i}\mathrm{E}[\hat{q}_{r/a}]$, hence

$$\overline{q}_i = \sum_{r\in i}\overline{q}_r[\sum_{a\in A}\lambda_a]. \tag{3}$$

For estimator $\hat{q}_i$ to have no bias, both (2) and (3) must hold for any value of a mean route flow $\overline{q}_r$: by identification, it must hold that

$$\forall r \in R_i\,,\ \sum_{a\in A}\lambda_a = 1\,. \tag{4}$$

Lemma 2, of Feasible Differential. *Let* $(\lambda_a)_{a\in A}$ *be a vector indexed by the arcs of G and let* ρ *be any real number. Then* $\sum_{a\in r}\lambda_a = \rho$ *for a set* R_i *of* $o-d$ *routes* r *if and only if there exists a field of node potentials* $\pi = [\pi(n) : n \in N] \in \Re^N$ *such that*

i. $\pi(o) = 0$,
ii. $\pi(d) = \rho$ *and*
iii. $\lambda_a = \pi(\mathrm{n}_a^-) - \pi(\mathrm{n}_a^+)$ *for every* $a \approx (\mathrm{n}_a^+, \mathrm{n}_a^-) \in A$.

Proof. Let us define $\lambda_r = \sum_{a\in r}\lambda_a$ for any route r and denote by $r \geq n$ [resp. $r < n$] the upstream subpath of r from its origin up to n [resp. the part of r strictly downstream of n]. Indeed, if such a network potential π exists, then λ satisfies trivially $\lambda_r = \rho$ for any $o-d$ route *r*. Conversely, suppose that such a λ exists. Define

$\pi(o) := 0$. For any node n in G distinct from o and such that there is an $o-d$ route *r* passing through *n*, define $\pi_r(n) := \lambda_{r \geq n}$: in fact $\pi_r(n)$ does not depend on *r* since for any other o-d route *r'* passing through *n*, from the constraints on λ we have that $\pi_r(n) + \lambda_{r \leq n} = \rho$ as well as $\pi_{r'}(n) + \lambda_{r' \leq n} = \rho$ but $(r \geq n) \bigcup (r' \leq n)$ makes an *o-d* route hence $\lambda_{r \geq n} + \lambda_{r' \leq n} = \rho$ which implies that $\lambda_{r \geq n} = \lambda_{r' \geq n} = \pi(n)$. Let us now check points i, ii, iii.

i. $\pi(o) = 0$ by definition.

ii. $\pi(d) = \rho$ since the sum of the λ_a on a route from o to d is equal to ρ.

iii. Take an arc $a \approx (n, n')$ traversed by an *o-d* route *r*: as $\rho = \lambda_{r \geq n} + \lambda_a + \lambda_{r \leq n'}$ with $\rho = \pi(d)$, knowing that $\lambda_{r \geq n} = \pi(n)$ and $\lambda_{r \leq n'} = \pi(d) - \pi(n')$, it holds that $\lambda_a = \pi(n') - \pi(n)$.

Actually Lemma 2 holds for the subnetwork $G_{od} = [N_{od}, A_{od}]$ defined by $N_{od} \equiv N_o^+ \bigcap N_d^-$ with node set N_o^+ comprised of all nodes *n* reachable from *o* by a positive $o-n$ route and node set N_d^- made up of all nodes *n* connected to *d* by a positive $n-d$ route, the arc set A_{od} being the restriction of A to the arcs with both head and tail in N_{od}. To ensure Lemma 2 to hold for the whole of A, let $\pi(n) := \pi(d)$ on $N_d^- \setminus N_o^+$, let $\pi(n) := \pi(o)$ on $N_o^+ \setminus N_d^-$ and let $\pi(n) := 0$ out of $N_o^+ \cup N_d^-$. Thus $\lambda_a = 0$ for $a \in A \setminus A_{od}$, which is consistent with $q_a = 0$ and $\sigma_a^2 = 0$ since no $o-d$ route passes through *a*.

The Lemma's name was devised after (*8*).

Bringing the definition of an optimal estimator together with Lemma 1, a linear combination $(\lambda_a)_{a \in A}$ is optimal if and only if it solves the following convex program:

$$\underset{\lambda \in \Re^A}{\text{Min}} \sum_{a \in A} \sigma_a^2 \lambda_a^2 \; ; \qquad \sum_{a \in r} \lambda_a = 1 \qquad \text{for every } o-d \text{ route } r. \tag{5}$$

In the present form, the set of constraints has an exponential description. On using Lemma 2 to reduce the number of constraints, the program above is equivalent to

$$\underset{\pi \in \Re^N}{\text{Min}} \sum_{(n,n') \in A} \sigma_{(n,n')}^2 [\pi(n') - \pi(n)]^2 \; ; \qquad \pi(o) = 0 \; ; \qquad \pi(d) = 1 \, , \tag{6}$$

or, in a more compact way,

$$\underset{\pi \in \Re^N}{\text{Min}} \left\| \mathrm{E}_\sigma \pi \right\|^2 ; \qquad \pi(o) = 0 \; ; \qquad \pi(d) = 1 \, . \tag{7}$$

To solve this convex minimization program subject to linear constraints, let us write the Lagrangean function $\mathcal{L} = \left\| \mathrm{E}_\sigma \pi \right\|^2 + \mu_o \pi(o) + \mu_d [\pi(d) - 1]$, in which "dual" variables μ_o and μ_d are associated to the potential constraints at node *o* and *d*, respectively. Differentiating the Lagrangian with respect to the variables $\pi(n)$ for $n \in N$ yields:

$$\frac{\partial \mathcal{L}}{\partial \pi(n)} = 2(\mathrm{E}_\sigma \mathrm{E}_\sigma^{\mathrm{t}} \pi)(n) \qquad \text{for } n \notin \{o, d\} \, , \tag{8}$$

$$\frac{\partial \mathcal{L}}{\partial \pi(o)} = 2(\mathrm{E}_\sigma \mathrm{E}_\sigma^{\mathrm{t}}\,\pi)(o) + \mu_o, \text{ and}$$

$$\frac{\partial \mathcal{L}}{\partial \pi(d)} = 2(\mathrm{E}_\sigma \mathrm{E}_\sigma^{\mathrm{t}}\,\pi)(d) + \mu_d\,.$$

As it is easy to define μ_o and μ_d so as to satisfy the above relations, the following conditions are necessary and sufficient for the optimality of a potential field π:

$$(\mathrm{E}_\sigma \mathrm{E}_\sigma^{\mathrm{t}}\,\pi)(n) = 0 \text{ for } n \notin \{o,d\}, \qquad (9)$$

$$\pi(d) = 1,$$

$$\pi(o) = 0,$$

which is exactly the claim in Theorem 1 since $\mathrm{E}_\sigma \mathrm{E}_\sigma^t = \mathrm{L}_{\sigma^2}$.

3.3 The issues of existence and uniqueness

The problem as stated in (7) is a quadratic minimization program with non negative coefficients: there exists an optimal solution if and only if there is a feasible solution, which can be derived on the basis of the following Lemma.

Lemma 3, Feasible Solution. *The program (9) admits a feasible solution if and only if o and d are positively connected and there is an o-d cut* $j = [S : N \setminus S]$ *made up of informative arcs with* $\sigma_a^2 < \infty$. *The associated feasible solution is* $\pi(n) := 0$ *at* $n \in S$ *and* $\pi(n) := 1$ *at* $n \in N \setminus S$.

Proof. If there is an *o-d* route which does not traverse any informative arc, then it is impossible to estimate the OD flow since that route flow cannot be estimated. A necessary and sufficient condition for no uninformative $o-d$ route to exist, is that an informative cut exists that separates *d* from *o*: it is sufficient because it prohibits the existence of an uninformative $o-d$ route, and it is necessary because if no such cut exists then either there is an uninformative $o-d$ route or there is no $o-d$ route, meaning an infeasible problem. The feasible solution associated to an informative cut *j* yields $\lambda_a := e_j(a)/\sigma_a^2$ along the cut and $\lambda_a := 0$ out of it.

About uniqueness, in general the linear system provided by Theorem 1 is non-singular, meaning that the optimal linear estimator is unique. More precisely, we have the following Proposition.

Proposition 1, Uniqueness of optimal estimator. *If all the* $\hat{q}_{ia}$ *have strictly positive variances, then the system (9) provided by the Theorem 1 is non-singular.*

Proof. Let us suppose that there are two optimal linear estimators $\hat{q}_i = \sum \lambda_a \hat{q}_{ia}$ and $\hat{q}'_i = \sum \lambda'_a \hat{q}_{ia}$, with $\mathrm{var}[\hat{q}_i] = \mathrm{var}[\hat{q}'_i]$. It holds that

$$\begin{aligned}\mathrm{var}[\hat{q}_i] &\le \mathrm{Var}[\tfrac{1}{2}\hat{q}_i + \tfrac{1}{2}\hat{q}'_i] = \textstyle\sum (\frac{\lambda_a + \lambda'_a}{2})^2\, \mathrm{var}[\hat{q}_{ia}] \\ &\le \tfrac{1}{2}(\textstyle\sum \lambda_a^2\, \mathrm{var}[\hat{q}_{ia}] + \sum \lambda_a'^2\, \mathrm{var}[\hat{q}_{ia}]) = \mathrm{var}[\hat{q}_i]\end{aligned}.$$

The second inequality stems from the fact that $(\lambda_a + \lambda'_a)^2 \le 2(\lambda_a^2 + \lambda_a'^2)$.

Hence, all inequalities are equalities, and $(\frac{\lambda_a+\lambda'_a}{2})^2 = \frac{1}{2}(\lambda_a^2+\lambda_a'^2)$ for every arc a such that $\mathrm{var}[\hat{q}_{ia}] \neq 0$, i.e. $\lambda_a = \lambda'_a$ for every arc a such that $\mathrm{var}[\hat{q}_{ia}] \neq 0$.

For instance, if there are several $o-d$ cuts with variance zero, the system is singular. Hereafter a related remark about an optimal linear estimator of null variance zero is given, which may simplify in some cases the construction of the estimator.

Proposition 2. *If the optimal linear estimator has null variance, then there is an optimal linear estimator associated to an $o-d$ cut.*

Proof. Take an optimal linear estimator of variance zero. Whenever $\lambda_a \neq 0$ we have $\mathrm{var}[\hat{q}_{ia}] = 0$. Since the λ_a come from a potential field π, there is a nonzero λ_a on any $o-d$ path r (directed or not). Hence, if we delete all arcs with $\lambda_a \neq 0$, o and d are in distinct connected components (connected in the sense of the underlying undirected graph). There is therefore an $o-d$ cut with $\lambda_a \neq 0$ for all its arcs, hence $\mathrm{var}[\hat{q}_{ia}] = 0$ for all its arcs, implying that the cut estimator has variance zero and is optimal.

3.4 Energy minimization problem in arc flows

In the linear system that characterizes the optimal node potentials, each node constraint formulated as $\partial\mathcal{L}/\partial\pi(n) = 0$ at node n can be interpreted as a flow conservation constraint at n. This leads us to design a problem of network flow, precisely a nonlinear distribution problem, which is equivalent to the node potential problem and enables one to deal with uninformative links in a straightforward way.

Let us restate system (8) as:

$$\frac{1}{2}\frac{\partial\mathcal{L}}{\partial\pi(n)} = (\mathrm{E}_\sigma \mathrm{E}_\sigma^t \pi)(n) + b_n = 0 \tag{10}$$

with $b_o = \mu_o/2$, $b_d = \mu_d/2$ and $b_n = 0 \ \ \forall n \in N \setminus \{o,d\}$.

As $\mathrm{E}_\sigma^t \pi = [\sigma_a(\pi(\mathrm{n}_a^+) - \pi(\mathrm{n}_a^-))]_{a\in A}$ and

$$\begin{aligned}(\mathrm{E}_\sigma \mathrm{E}_\sigma^t \pi)(n) &= \textstyle\sum_{a\in A} e(n,a)\sigma_a^2(\pi(\mathrm{n}_a^+) - \pi(\mathrm{n}_a^-)) \\ &= \textstyle\sum_{a\in A_n^+} \sigma_a^2(\pi(n) - \pi(\mathrm{n}_a^-)) - \sum_{a\in A_n^-} \sigma_a^2(\pi(\mathrm{n}_a^+) - \pi(n))\end{aligned}$$

the system is equivalent to

$$\forall n\in N \quad \textstyle\sum_{a\in A_n^+} \sigma_a^2[\pi(\mathrm{n}_a^-) - \pi(n)] - \sum_{a\in A_n^-} \sigma_a^2[\pi(n) - \pi(\mathrm{n}_a^+)] = b_n$$

On replacing $\sigma_a^2[\pi(\mathrm{n}_a^-) - \pi(\mathrm{n}_a^+)]$ by an arc flow x_a, we get that

$$\forall n\in N \quad \textstyle\sum_{a\in A_n^+} x_a - \sum_{a\in A_n^-} x_a = b_n, \tag{11}$$

or in matrix notation $\mathrm{E}x = b$,

which states the conservation of flow at node *n*, or more precisely: That at node *n* the network flow $x = [x_a]_{a\in A}$ receives a net inflow b_n from the outside. At nodes distinct from o and d the net inflow is zero, meaning flow conservation. At o the network imports b_o whereas at *d* the network imports b_d : then it must hold that $b_d = -b_o$.

We also characterize flow x by the requirement that there exists a potential vector π such that $x_a = \sigma_a^2[\pi(\mathrm{n}_a^-) - \pi(\mathrm{n}_a^+)]$. Let $R_a = 1/\sigma_a^2$ if $\sigma_a^2 > 0$ or 0 if $\sigma_a^2 = \infty$ or $R_{\max}$ otherwise – an arbitrarily large value to cope with the case of zero σ_a^2 which means a deterministic knowledge of link OD flow q_{ia}, and let A^* be the subset of informative links with $\sigma_a^2 < \infty$: it holds that

$$\forall a \in A^* \quad \pi(\mathrm{n}_a^-) - \pi(\mathrm{n}_a^+) = R_a x_a \tag{12}$$

An uninformative arc has infinite variance so x_a is finite only if $\pi(\mathrm{n}_a^-) = \pi(\mathrm{n}_a^+)$, which further requires $R_a x_a$ to be zero. This is compatible with (12) applied to $a \in A \setminus A^*$ with $R_a = 0$, so from now on the distinction between A^* and A can be ignored. In matrix notation, letting $\mathrm{R}_A = \mathrm{Diag}[R_a : a \in A]$, the extension of (12) to A is restated as

$$-\mathrm{E}^t \pi = \mathrm{R}_A x \tag{13}$$

The set of conditions (11), (13) amounts to the primal-dual optimality conditions of the following optimization program:

$$\min_x \mathrm{J}(x) \equiv \tfrac{1}{2} x^t \mathrm{R}_A x = \tfrac{1}{2} \textstyle\sum_{a \in A} R_a x_a^2 \quad \text{s.t. } \mathrm{E}x = b \tag{14}$$

since those primal-dual conditions are made up of (11) together with the zeroing of the Lagrangian derivatives $\partial \mathcal{L}_\mathrm{J} / \partial x_a = R_a x_a + \pi(\mathrm{n}_a^+) - \pi(\mathrm{n}_a^-)$, in which $\pi(n)$ now stands for the dual variable associated to the n-th constraint $(\mathrm{E}x)(n) = b_n$.

Thus the network flow is an optimal solution to the problem of minimizing some kind of electrical energy. The arc resistance $R_a = 1/\sigma_a^2$ is the reciprocal of the link variance, hence it is homogeneous to the precision (or accuracy) in a Bayesian mixture. We can also think of x as an imprecision flow, in consistency with

$$\text{Imprecision} \approx \text{Variance} \approx \sigma^2 (\Delta\pi)^2 \approx \sigma^2(\Delta\pi) \approx x \quad \text{since } \Delta\pi \text{ is a number}$$

To sum up our discussion, let us state:

Theorem 2, Optimal imprecision flow. *On the oriented graph $G = [N, A]$ with node-arc incidence matrix E, consider the diagonal matrix R_A of n-th term $R_a = 1/\sigma_a^2$ and the inflow vector b with components $+1$ at node o, -1 at node d and 0 elsewhere. The network flow x that minimizes $\mathrm{J}(x) = \frac{1}{2} x^t \mathrm{R}_A x$ subject to $\mathrm{E}x = b$ yields node potentials $\pi(n)$ and arc tensions $v_a = x_a / \sigma_a^2$. Letting $\lambda_a := v_a / (\pi(d) - \pi(o))$, the coefficients $(\lambda_a)_{a \in A}$ provide an optimal unbiased linear estimator $\hat{q}_i := \sum_{a \in A} \lambda_a \hat{q}_{ia}$ of the network $o-d$ flow $\overline{q}_i$ on the basis of the independent link origin-destination estimators $\hat{q}_{ia}$ of variance σ_a^2.*

A range of algorithms were designed to solve the electrical network problem (14), including methods analogous to traffic assignment algorithms since electrical equilibrium closely resembles traffic equilibrium – see e.g. (*9, 10*).

3.5 The case of dependent local estimators

If the link estimators $\hat{q}_{ia}$ are independent then their covariance satisfies $\mathrm{cov}[\hat{q}_{ia}, \hat{q}_{ib}] = 0$ if $a \neq b$. However some dependencies may arise, yielding non zero covariance that must be taken into account. It is straightforward to extend the node potential problem to that case: denoting by $v_{ab} := \mathrm{cov}[\hat{q}_{ia}, \hat{q}_{ib}]$ and knowing that $\sigma_a^2 = \mathrm{cov}[\hat{q}_{ia}, \hat{q}_{ia}]$, the variance of a network estimator becomes

$$\mathrm{var}[\hat{q}_i] = \sum_{a,b \in A} \lambda_a \lambda_b v_{ab} . \tag{15}$$

With respect to node potentials that are still in order since an unbiased estimator is required, then

$$\mathrm{G}(\pi) := \mathrm{var}[\hat{q}_i] = \sum_{a,b \in A} v_{ab} [\pi(\mathrm{n}_a^-) - \pi(\mathrm{n}_a^+)][\pi(\mathrm{n}_b^-) - \pi(\mathrm{n}_b^+)] \tag{16}$$

The problem of variance minimization still has first-order optimality conditions that are linear in the $\pi(n)$ variables, since

$$\frac{\partial \mathrm{G}}{\partial \pi(n)} = 2\sum_{a \in A} e(n,a) \{ \sigma_a^2 [\pi(\mathrm{n}_a^-) - \pi(\mathrm{n}_a^+)] + \sum_{b \neq a} v_{ab} [\pi(\mathrm{n}_b^-) - \pi(\mathrm{n}_b^+)] \} \tag{17}$$

The linear system of zeroing the Lagrangian derivatives that combine $\partial \mathrm{G} / \partial \pi(n)$ with the potential constraints at nodes *o* and *d*, is still easy to solve. The feasibility and existence condition is the same as in the independent case because the essential requirement is to have an informative cut separating *d* from *o*, be it made up of independent or dependent local estimators.

4. OPTIMAL ESTIMATOR IN PRACTICE

If a measure $\hat{q}_{ia}$ is available for every arc a used by the $o-d$ flow, the best linear estimator is easily computed from the system in Theorem 1 through Gaussian elimination. In this Section, we will see how to deal with arcs for which no information is available and how to check if there are enough measures $\hat{q}_{ia}$ to derive a network estimator $\hat{q}_i$. This can provide some indication in the design of an observation scheme. Several examples of the method will be given. Lastly, we shall analyze the computational complexity of the algorithm.

4.1 What if no information is available for some arcs?

Actually, the answer is easy. In the convex program in the proof of Theorem 1, the index set of the sum $\sum_a \sigma_a^2 \lambda_a^2$ is restricted to those arcs *a* for which a measure $\hat{q}_{ia}$ is available of finite variance $\sigma_a^2 < \infty$. Now, for any arc $a \approx (n, n')$ for which no information is available i.e. $\sigma_a^2 = \infty$, let us define $\lambda_a := 0$ and take the convention that $0.\infty = 0$. This does not alter the optimal value of the convex program if it is feasible, and now the sum of the λ_a along any $o-d$ route is well-defined and equal to 1. The remainder of the proof of Theorem 1 becomes valid and yields that $\pi(n) = \pi(n')$ wherever no local information is available on arc (n, n'). This means that we could have contracted any such arc without modifying the solution. The *contraction* of an arc is a classical operation in graph theory (*11*). Recalling that a *loop* is an arc of which the endpoints are identical, to *contract* an arc means:

- To unify its two endpoints into a single node;
- To delete all loops that may have appeared.

The method for finding the best linear estimator can be summarized as follows:

Contract all arcs for which no information is available. Then solve the system in Theorem 1 for the contracted graph, in which all the arcs are informative for the o-d pair. Lastly, coming back to the original network, assign to each original node the potential of the associated unified node in the contracted network, and to each original arc the tension λ_a that stems from the potentials of its endpoints.

Note that the sequence of contractions is easily computed, for instance on the basis of a forward-star representation of the network. The contracted graph yields a linear system of reduced dimension, which allows for efficient solution.

4.2 Some indication on the design of the measures

According to the previous Subsection if, after all possible contractions, the nodes o and d have been unified, then no *o-d* flow can be estimated from the available measures. An alternative test is:

Delete all arcs for which a measure is available. An estimator of the $o-d$ flow can be derived if and only if o and d do not remain in the same connected component.

Note that it is easy to test if two nodes are in the same connected component (with depth-first-search for instance).

Moreover, loops with available measure that are deleted during the contraction steps are useless: the measures on such loops are lost. Hence,

To avoid useless measurement, the set of measured arcs must be the union of undirected $o-d$ cuts.

4.3 Examples

Two arcs in series. Consider the graph with two arcs in series: (o,n) and (n,d) as in Figure 3a. Suppose that the two measures $\hat{q}_{(o,n)}$ and $\hat{q}_{(n,d)}$ are available and have the same standard deviation $\sigma_a = 1$. Thus

$$\mathrm{L}_{\sigma^2} = \begin{array}{c} \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \begin{array}{l} o \\ n \\ d \end{array} \\ \begin{array}{ccc} o & \quad n & \quad d \end{array} \end{array}$$

Let us apply Theorem 1: $\pi(o) = 0$, $\pi(d) = 1$ and $-\pi(o) + 2\pi(n) - \pi(d) = 0$. Thus, $\pi(n) = \frac{1}{2}$. The optimal coefficients for the arcs (o,n) and (n,d) are both equal to $\frac{1}{2}$. The optimal estimator $\hat{q}$ is equal to $\frac{1}{2}\hat{q}_{(o,n)} + \frac{1}{2}\hat{q}_{(n,d)}$, in accordance with the intuition that the serial combination of observations amounts to pooling them with specific weights.

Two arcs in parallel. Consider the graph with two arcs a and a' in parallel as in Figure 3b. Suppose that the two measures $\hat{q}_a$ and $\hat{q}_{a'}$ have the same standard deviation $\sigma_a = \sigma_{a'} = 1$. According to Theorem 1, the optimal coefficients are both equal to $\pi(d) - \pi(o) = 1$. The optimal estimator $\hat{q}$ is then equal to $\hat{q}_a + \hat{q}_{a'}$, in accordance with the intuition that the parallel decomposition of flow corresponds to a summation.

The realistic example of Section 2. Consider the realistic example in Section 2. Contraction of the uninformative arcs yields the graph depicted in Figure 4. This graph has three nodes: o, n and d and five arcs A, B, C, D, E. Denoting also by A, B, C, D, E the variances of the corresponding arcs, we obtain that

$$\mathrm{L}_{\sigma^2} = \begin{array}{cc} \begin{pmatrix} A+B & -A & -B \\ -A & A+C+D+E & -C-D-E \\ -B & -C-D-E & B+C+D+E \end{pmatrix} & \begin{array}{c} o \\ n \\ d \end{array} \\ \begin{array}{ccc} \quad o \quad & \qquad n \qquad & \qquad d \qquad \end{array} & \end{array}$$

Application of Theorem 1 yields that $\pi(d) = 1$ and also $-A\pi(o) + (A+C+D+E)\pi(n) - (C+D+E)\pi(d) = 0$. This leads to

$$\pi(n) = \frac{C+D+E}{A+C+D+E} \text{ at the Cognac node.}$$

Hence $\lambda_A = \pi(n) \approx 0.41$, $\lambda_B = 1$, $\lambda_C = \lambda_D = \lambda_E = 1 - \pi(n) \approx 0.59$.

Thus the optimal estimator of the OD flow is

$$\sum_a \lambda_a \hat{q}_{ia} \approx 0.59(q_{iC} + q_{iD} + q_{iE}) + 0.41 q_{iA} + 1.0 q_{iB} \approx 182 \text{ veh/day,}$$

while its variance amounts to $\sum_a \lambda_a^2 \sigma_a^2 \approx 0.41^2(C+D+E) + 0.59^2 A + 1.0^2 B \approx 923$, yielding $\mathrm{SE} \approx 30{,}4$ veh/day.

This is consistent with, and simpler than, the by-hand computation in Section 2.

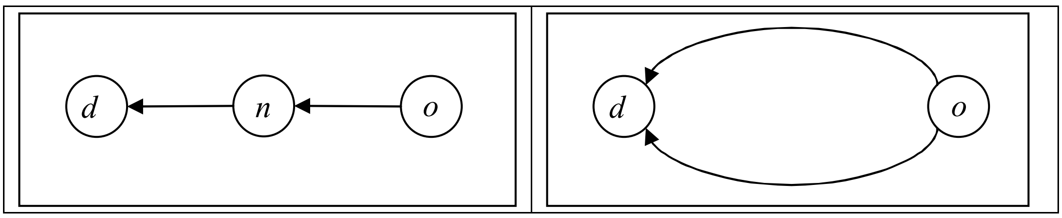

FIGURE 3 (a) Two arcs in series; (b) Two arcs in parallel

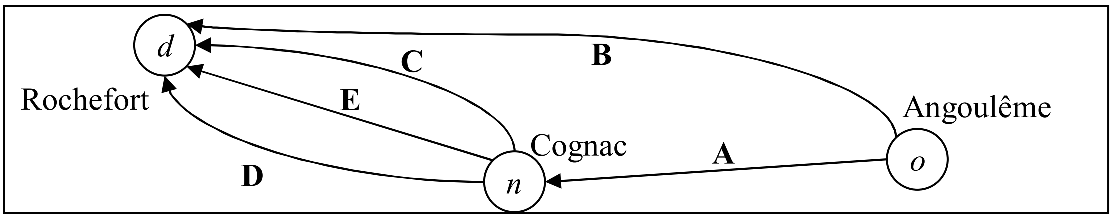

FIGURE 4 Contracted graph of the Angoulême-Rochefort OD pair

4.4 Computation scheme and its complexity

Let $\overline{n}$ (resp. $\overline{a}$) denote the number of nodes (resp. arcs) in the original network, and $\bar{t}$ be the size of the set T of informative links.

The graph contraction to keep only the informative links and to delete the loops requires a number $O(\overline{n}+\overline{a})$ of elementary operations. To prove this claim an efficient algorithm is provided hereafter, since a naïve implementation would lead to a complexity of $O(\overline{n}\,\overline{a})$. The algorithm includes the following steps:

- **Initialization**. Assume a forward-star representation of the original network with node field $\mathrm{FO}[n]$ to indicate the first position of an outgoing arc $a \in A_n^+$ in the array of arcs, and arc field $\mathrm{HO}[a]$ to indicate the head node of arc a.
- **Incoming arcs**. Associate to each node its incoming arcs by building a backward-star representation of the original network, with node field $\mathrm{FI}[n]$ to indicate the first position of an incoming arc $a \in A_n^-$ in the backward array of arcs, and arc field $\mathrm{TI}[a]$ to indicate the tail node of arc a. An auxiliary field $\mathrm{NI}[n]$ is required to count the arcs coming in n. This works as follows:
    - *For* $n := 1$ *to* $\overline{n}$ *let* $\mathrm{NI}[n] := 0$*;*
    - *For* $n := 1$ *to* $\overline{n}$*, for* $a := \mathrm{FO}[n]$ *to* $\mathrm{FO}[n+1]-1$*,* $\mathrm{NI}[\mathrm{HO}[a]] := \mathrm{NI}[\mathrm{HO}[a]]+1$*;*
    - *Let* $\mathrm{FI}[1] := 1$*;*
    - *For* $n := 1$ *to* $\overline{n}$ *let* $\mathrm{FI}[n+1] := \mathrm{FI}[n] + \mathrm{NI}[n]$ *and* $\mathrm{NI}[n] := 0$*;*
    - *For* $n := 1$ *to* $\overline{n}$*, for* $a := \mathrm{FO}[n]$ *to* $\mathrm{FO}[n+1]-1$*, Let* $m := \mathrm{HO}[a]$*,* $\mathrm{TI}[\mathrm{FI}[m]+\mathrm{NI}[m]] := n$ *and* $\mathrm{NI}[m] := \mathrm{NI}[m]+1$.
- **Connected components**. Assign each node n to its connected component in network $[N, A \setminus T]$, i.e. the subset of all nodes m such that there is a path from n to m made up of arcs in $A \setminus T$ whatever the arc direction. The method to discover the connected component is to mark the nodes progressively, only one time each: on the first time that an arc connection reveals that the node belongs to the component under construction, the node is included in a chained list to wait for treatment, which consists in searching its own connections to other nodes. To that end the following fields are required: field $\mathrm{CN}[n]$ to relate node n to its component, which is referred to by its minimum node index; component index $\mathrm{CI}[n]$ to relate node n to its component index $n°$ in $\{1,..\tilde{n}\}$ where $\tilde{n}$ stands for the number of connected components; and reverse component $\mathrm{rCI}[n°]$ to relate component index $n°$ to its node of minimum node index; and also a "next node" field $\mathrm{NN}[n]$ to chain node n to the next node in its connected component, or to contain 0 if n is the last node in that list. Here are the detailed operations:
    - *Let* $\tilde{n} := 0$*;*
    - *For* $n := 1$ *to* $\overline{n}$ *let* $\mathrm{CN}[n] := n$*,* $\mathrm{NN}[n] := 0$*,* $\mathrm{CI}[n] := 0$ *and* $\mathrm{rCI}[n] := 0$*;*
    - *For* $n := 1$ *to* $\overline{n}$*, if* $n = \mathrm{CN}[n]$ *then:*
        - *Let* $\tilde{n} := \tilde{n}+1$*,* $\mathrm{CI}[n] := \tilde{n}$ *and* $\mathrm{rCI}[\tilde{n}] := n$*;*
        - *Let* $f := n$ *and* $\ell := n$*; // pointers on first and last elements in list;*

- *Repeat:*
  - *Let* $\mathrm{CN}[f] := n$ *and* $\mathrm{CI}[f] := \tilde{n}$ *;*
  - *for* $a := \mathrm{FO}[f]$ *to* $\mathrm{FO}[f+1]-1$*, if (*$a \notin T$*) and (*$\mathrm{NN}[\mathrm{HO}[a]] = 0$*) then let* $\mathrm{NN}[\ell] := \mathrm{HO}[a]$ *and* $\ell := \mathrm{HO}[a]$*;*
  - *for* $a := \mathrm{FI}[f]$ *to* $\mathrm{FI}[f+1]-1$*, if (*$a \notin T$*) and (*$\mathrm{NN}[\mathrm{TI}[a]] = 0$*) then let* $\mathrm{NN}[\ell] := \mathrm{TI}[a]$ *and* $\ell := \mathrm{TI}[a]$*;*
  - *Let* $f := \mathrm{NN}[f]$*;*

  Until ($f = 0$*).*

- **Contracted network**. Build a forward-star representation of the contracted network, excluding loops, by using node array $\mathrm{FC}[n°]$ of outgoing informative arcs and arc array $\mathrm{HC}[a°]$ of head node. An arc field $\mathrm{OA}[a°]$ relates each arc $a°$ in the contracted network to its index a in the original network. An auxiliary node field $\mathrm{NA}[n°]$ is required to count the number of informative arcs that go out of $n°$. This works as follows:
  - *Define* $\tilde{a}$ *the number of contracted arcs and initialize its value at zero.*
  - *For* $n := 1$ *to* $\bar{n}$*, for* $a := \mathrm{FO}[n]$ *to* $\mathrm{FO}[n+1]-1$*, if (*$a \in T$*) and (*$\mathrm{CN}[\mathrm{HO}[a]] \neq \mathrm{CN}[n]$*) then let* $\mathrm{NA}[\mathrm{CI}[n]] := \mathrm{NA}[\mathrm{CI}[n]]+1$ *and* $\tilde{a} := \tilde{a}+1$*;*
  - *Let* $\mathrm{FC}[1] := 1$*;*
  - *For* $n° := 1$ *to* $\tilde{n}$ *do* $\mathrm{FC}[n°+1] := \mathrm{FC}[n°] + \mathrm{NA}[n°]$ *and* $\mathrm{NA}[n°] := 0$*;*
  - *For* $n := 1$ *to* $\bar{n}$*, for* $a := \mathrm{FO}[n]$ *to* $\mathrm{FO}[n+1]-1$*, if (*$a \in T$*) and (*$\mathrm{CN}[\mathrm{HO}[a]] \neq \mathrm{CN}[n]$*) then let* $\mathrm{NA}[\mathrm{CI}[n]] := \mathrm{NA}[\mathrm{CI}[n]]+1$*,* $b := \mathrm{FC}[\mathrm{CI}[n]] + \mathrm{NA}[\mathrm{CI}[n]]$*,* $\mathrm{HC}[b] := \mathrm{CI}[\mathrm{HO}[a]]$ *and* $\mathrm{OA}[b] := a$.

As each step Initialization, Incoming arcs, Connected components and Contracted network has a computational complexity of $O(\bar{n}+\bar{a})$, so does the sequence of them.

Now, letting $\bar{z}$ denote the number of traffic demand zones, there are $\tilde{z}$ contracted zones with respect to the contracted network, yielding at most $\tilde{z}^2 \le \bar{z}^2$ OD pairs.

For any contracted OD pair, the computation of the optimal network estimator is performed along the following steps:

- customization of the local information, yielding the link OD flow $\hat{q}_{ia}$ and variance σ_a^2 (in fact $\sigma_a^{(i)2}$) by contracted arc a for that OD pair, in $O(\tilde{t})$ operations.
- Building the customized Laplacian matrix $L_{\sigma^2}^{(i)}$, an $O(\tilde{a}) = O(\tilde{t})$ task since the matrix is made up of zeroes save for the accumulation of σ_a^2 in cells $(\mathrm{n}_a^+, \mathrm{n}_a^-)$ and $(\mathrm{n}_a^-, \mathrm{n}_a^+)$ and in the diagonal line.
- Solving $L_{\sigma^2}^{(i)} \pi^{(i)}$ equal to a simple vector in the right-hand side: by using Gaussian elimination this amounts to a $O(\tilde{n}^3)$ task.
- Deriving the arc coefficients $\lambda_a^{(i)}$ from the node potentials $\pi_n^{(i)}$, an $O(\tilde{a})$ task.

Overall, the computational complexity of the OD treatment amounts to $O(\tilde{n}^3)$, making the whole computation of an "optimally estimated" trip matrix an $O(\tilde{z}^2\tilde{n}^3 + \overline{n} + \overline{a})$ task.

5. CONCLUSION

A systematic, efficient yet relatively simple method was provided to estimate an OD flow at the network level on the basis of link OD information on a subset of network links. The problem of optimal estimation is stated as the estimator of minimal variance among the linear combinations of link-based estimators that are unbiased. This problem is endowed with nice graph-theoretic properties: notably, the combinatorics are reduced to a linear system in node potentials, or equivalently (in the independent case) to a linear system in arc flows.

Our model does not require any assumption about route choice proportions, contrarily to entropy-based models (*12*) and related statistical approaches (*13*). Based on this fact, future work could be targeted at the joint estimation of OD flows and route choice behaviour, on the basis of additional information about route attributes and of a behavioural model of traffic assignment to road networks. Among the topics that also deserve further research, let us mention: the use of additional link counts deprived of OD link surveys; the consideration of counts between link pairs, particularly so when a cordon trip matrix is available; the treatment of biased local estimators.

6. REFERENCES

(1) Beckmann, M., McGuire, C.B., & Winsten, C.B. (1956). *Studies in the Economics of Transportation*. Yale University Press, New Haven, Connecticut.

(2) Sheffi Y (1984) *Urban Transportation Networks.* Prentice Hall, Englewood Cliffs, New Jersey.

(3) Ortuzar, J. de Dioz, Willumsen, L.G. (2004). *Modeling Transport*. Wiley, London. First edition, 1990.

(4) Taylor, M.A.P., Young, W. (1988). *Traffic Analysis: new technology and new solutions*. Hargreen, Victoria, Australie.

(5) Leurent F (2002) *Estimer un modèle de distribution spatiale avec des enquêtes en-route.* In Aron M, Boillot F and Lebacque JP (eds) Modélisation du trafic, actes du groupe de travail 1999. Actes INRETS #83, pp. 85-102. INRETS, France.

(6) Wonnacott, T.H., & Wonnacott, R.J. (1991). *Statistique*. Economica, Paris.

(7) Kuwahara, M., Sullivan, E.C. (1987) Estimating origin-destination matrices from roadside survey data. *Transportation Research B*, **21**, 233-248.

(8) Rockafellar RT (1984) *Network Flows and Monotropic Optimization*. John Wiley and Sons, New York.

(9) Wellstead, P.E. (1979) Introduction to Physical System Modeling. *Academic Press*.

(10) Fenton, N., Hill, G. (1993). Systems construction and analysis. *McGraw-Hill International*.

(11) Diestel R (2006). Graph Theory. *Springer*.

(12) Van Zuylen, H.J., & Willumsen, L.G. (1980). The most likely trip matrix estimated from traffic counts. *Transportation Research B*, **14**, 281-293.

(13) Cascetta, E., Nguyen, S. (1988). A unified framework for estimating or updating origin/destination matrices from traffic counts. *Transportation Research B*, **22**/6, 437-455.